

# SEQUENCE SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 03:18:06 ; Search time 8804 Seconds  
(without alignments)  
11070.431 Million cell updates/sec

Title: US-10-089-514-1  
Perfect score: 2061  
Sequence: 1 atgcgcacgcttctgatcga.....ccgtggcgggcgcccgatga 2061  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2061	100.0	2061	6 BD178313	GenBank BD178313 Transform
2	2061	100.0	2061	6 BD093914	" BD093914 Transform
3	2061	100.0	5251	1 AB116234	" AB116234 Streptomy
4	1773.8	86.1	2660	1 SVU21728	GenBank U21728 Streptomyce
5	1773.8	86.1	14159	1 AF262220	AF262220 Streptomy
6	820	39.8	2220	6 AR198361	AR198361 Sequence
7	820	39.8	4740	1 SPU60417	U60417 Streptomyce
8	773.4	37.5	4607	1 STMPABA	M93058 Streptomyce
9	773	37.5	39314	1 SGR300302	AJ300302 Streptomy
10	769.8	37.4	138203	1 AY310323	AY310323 Streptomy

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 03:16:03 ; Search time 1005 Seconds  
(without alignments)  
10765.231 Million cell updates/sec

Title: US-10-089-514-1  
Perfect score: 2061  
Sequence: 1 atgcgcacgcttctgatcga.....ccgtggcgggcgcccgatga 2061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2061	100.0	2061	4 AAF86409	wo 01/23542 Aaf86409 4-amino-4
2	2061	100.0	2061	8 AAL50180	wo 02/77244 Aal50180 S venezue
3	1773.8	86.1	3305	8 ABZ69799	Abz69799 Plasmid p
4	1772.2	86.0	12391	8 ABZ69798	Abz69798 Plasmid p
5	459	22.3	4496	2 AAT58553	Aat58553 Streptomy

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 04:42:01 ; Search time 181 Seconds  
(without alignments)  
8093.568 Million cell updates/sec

Title: US-10-089-514-1  
Perfect score: 2061  
Sequence: 1 atgcgcacgcttctgatcga.....ccgtggcgggcgcccgatga 2061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query				Description
	No.	Score	Match	Length	DB	ID			
	1	820	39.8	2220	3	US-08-765-907A-14			Sequence 14, Appl
	2	459	22.3	4496	3	US-08-765-907A-6			Sequence 6, Appli
	3	365.4	17.7	2888	3	US-08-765-907A-1			Sequence 1, Appli
	4	272.2	13.2	645	2	US-08-403-852D-9			Sequence 9, Appli
	5	272.2	13.2	645	3	US-08-510-646B-9			Sequence 9, Appli
	6	272.2	13.2	645	3	US-09-231-818-9			Sequence 9, Appli
	7	272.2	13.2	645	4	US-09-635-359B-9			Sequence 9, Appli
c	8	268.4	13.0	1101	4	US-09-252-991A-13413			Sequence 13413, A
	9	268.4	13.0	1431	4	US-09-252-991A-13733			Sequence 13733, A
	10	260	12.6	816	4	US-09-252-991A-13943			Sequence 13943, A
	11	197.8	9.6	1506	4	US-09-252-991A-811			Sequence 811, App

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 04:52:43 ; Search time 1044 Seconds  
(without alignments)  
10667.498 Million cell updates/sec

Title: US-10-089-514-1  
Perfect score: 2061  
Sequence: 1 atgcgcacgcttctgatcga.....ccgtggcgggcgcccgatga 2061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

instant application NOT published

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2061	100.0	2061	18	US-10-472-587-1	Sequence 1, Appli
2	1773.8	86.1	3305	14	US-10-126-927-68	Sequence 68, Appl
3	1772.2	86.0	12391	14	US-10-126-927-67	Sequence 67, Appl
4	606.2	29.4	2196	15	US-10-156-761-1167	Sequence 1167, Ap
c 5	606.2	29.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
6	301.4	14.6	9025608	15	US-10-156-761-1	Sequence 1, Appli
7	280.8	13.6	1479	15	US-10-156-761-6148	Sequence 6148, Ap
8	270	13.1	1362	16	US-10-282-122A-30135	Sequence 30135, A
9	266.8	12.9	1860	9	US-09-738-626-1111	Sequence 1111, Ap
10	266.8	12.9	3309400	9	US-09-738-626-1	Sequence 1, Appli
c 11	234.8	11.4	42000	17	US-10-159-257A-1	Sequence 1, Appli
c 12	234.8	11.4	63158	15	US-10-292-198-1	Sequence 1, Appli
13	231.6	11.2	1338	16	US-10-282-122A-32104	Sequence 32104, A
14	224.2	10.9	2523	16	US-10-282-122A-17825	Sequence 17825, A
15	220.4	10.7	1257	16	US-10-282-122A-33856	Sequence 33856, A
16	210.2	10.2	2160	14	US-10-138-927-76	Sequence 76, Appl
17	210.2	10.2	2160	14	US-10-138-927-83	Sequence 83, Appl
18	210.2	10.2	2160	15	US-10-430-011-76	Sequence 76, Appl
19	210.2	10.2	2160	15	US-10-430-011-83	Sequence 83, Appl
20	210.2	10.2	2160	15	US-10-430-011-126	Sequence 126, App
21	201.2	9.8	2199	15	US-10-430-011-122	Sequence 122, App
22	198.2	9.6	1353	16	US-10-282-122A-23903	Sequence 23903, A
23	196.4	9.5	636	15	US-10-156-761-4319	Sequence 4319, Ap
24	192.8	9.4	2256646	17	US-10-470-565-1	Sequence 1, Appli
25	190.4	9.2	1323	16	US-10-282-122A-19610	Sequence 19610, A
26	189.6	9.2	2190	14	US-10-138-927-75	Sequence 75, Appl
27	189.6	9.2	2190	15	US-10-430-011-75	Sequence 75, Appl
28	186	9.0	661	18	US-10-425-115-81629	Sequence 81629, A
29	182.2	8.8	2166	15	US-10-430-011-127	Sequence 127, App
30	180.8	8.8	2202	15	US-10-430-011-121	Sequence 121, App
31	178.6	8.7	65140	17	US-10-203-295-1	Sequence 1, Appli
32	178.6	8.7	125401	17	US-10-203-295-35	Sequence 35, Appl
33	173.8	8.4	2190	14	US-10-138-927-93	Sequence 93, Appl
34	173.8	8.4	2190	15	US-10-430-011-93	Sequence 93, Appl
35	173.6	8.4	1119	15	US-10-156-761-6781	Sequence 6781, Ap
36	170.6	8.3	2190	14	US-10-138-927-1	Sequence 1, Appli
37	170.6	8.3	2190	14	US-10-138-927-84	Sequence 84, Appl
38	170.6	8.3	2190	14	US-10-138-927-85	Sequence 85, Appl
39	170.6	8.3	2190	14	US-10-138-927-86	Sequence 86, Appl
40	170.6	8.3	2190	14	US-10-138-927-87	Sequence 87, Appl
41	170.6	8.3	2190	14	US-10-138-927-88	Sequence 88, Appl
42	170.6	8.3	2190	14	US-10-138-927-92	Sequence 92, Appl
43	170.6	8.3	2190	15	US-10-430-011-1	Sequence 1, Appli
44	170.6	8.3	2190	15	US-10-430-011-84	Sequence 84, Appl
45	170.6	8.3	2190	15	US-10-430-011-85	Sequence 85, Appl

#### ALIGNMENTS

RESULT 1

US-10-472-587-1

; Sequence 1, Application US/10472587

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 04:40:31 ; Search time 6560 Seconds  
(without alignments)  
11448.519 Million cell updates/sec

Title: US-10-089-514-1  
Perfect score: 2061  
Sequence: 1 atgcgcacgcttctgatcga.....ccgtggcgggcgccgatga 2061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	216.6	10.5	529	6	CA003778	CA003778 HS15I14r
2	212.8	10.3	845	8	BZ570522	BZ570522 msh2_1411
3	205.8	10.0	653	7	CF863831	CF863831 psZS008xG
4	195.6	9.5	1043	8	BZ548859	BZ548859 pacs1-60_
c 5	178	8.6	523	8	BZ412600	BZ412600 OGAAG11TM
6	178	8.6	631	8	BZ412593	BZ412593 OGAAG11TC
7	178	8.6	631	8	BZ537360	BZ537360 OGAGD95TC
c 8	178	8.6	810	9	CG170785	CG170785 PUIIR58TD